Plant Biotechnology Journal

Plant Biotechnology Journal (2008) **6**, pp. 832–842 doi: 10.1111/j.1467-7652.2008.00364.x

Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation

Monica A. Schmidt[1,2] and Eliot M. Herman[1,*]

[1]*Plant Genetics Research Unit USDA/ARS, Donald Danforth Plant Science Center, 975 N. Warson Road, St. Louis, MO 63132, USA*
[2]*Donald Danforth Plant Science Center, 975 N. Warson Road, St. Louis, MO 63132, USA*

Received 15 March 2008;
revised 21 June 2008;
accepted 23 June 2008.
**Correspondence* (fax 00 1 314587-1392;
e-mail eherman@danforthcenter.org)

Keywords: conglycinin, glycinin, green fluorescent protein, protein body, seed, soybean, transgenic.

Summary

Seeds possess a high intrinsic capacity for protein production that makes them a desirable bioreactor platform for the manufacture of transgenic products. One strategy to enhance foreign protein production involves exchanging the capacity to produce intrinsic proteins for the capacity to produce a high level of foreign proteins. Suppression of the α/α′ subunit of β-conglycinin storage protein synthesis in soybean has been shown previously to result in an increase in the accumulation of the glycinin storage protein, some of which is sequestered as proglycinin into *de novo* endoplasmic reticulum (ER)-derived protein bodies. The exchange of glycinin for conglycinin is quantitative, with the remodelled soybeans possessing a normal protein content with an altered proteome. The green fluorescent protein (GFP)-kdel reporter was transferred in a construct using the glycinin promoter and terminator to mimic glycinin gene expression. When expressed in soybean seeds, GFP-kdel accreted to form ER-derived protein bodies. The introgression of GFP-kdel into the α/α′ subunit of the β-conglycinin suppression background resulted in a fourfold enhancement of GFP-kdel accumulation to > 7% (w/w) of the total protein in soybean seeds. The resulting seeds accumulated a single population of ER membrane-bound protein bodies that contained both GFP-kdel and glycinin. Thus, the collateral proteome rebalancing that occurs with the suppression of intrinsic proteins in soybean can be exploited to produce an enhanced level of foreign proteins.

Introduction

The use of plants as protein bioreactor platforms has broad potential, but little current commercial application, largely because of the low protein production yields obtained in plants compared with those produced in microbial expression systems. Seeds are an ideal plant platform for the cost-efficient production of foreign proteins, as they are already optimized for maximum protein synthesis and accumulation. The difficulty in engineering seeds as a bioreactor platform is that they are developmentally determinant and evolved in such a manner that the maturation of each seed is essentially identical. Seeds have evolved to store triglycerides, non-structural reserve carbohydrates and protein at maximum density within storage cells, leaving little cellular space to add additional products resulting from transgene expression. Within its developmental programme, the seed exhibits a limited degree of storage substance plasticity, and rebalances storage protein content via nutrient availability, such as sulphur (for examples, see Beach *et al*., 1985; Hirai *et al*., 1995; Hagan *et al*., 2003; for a review, see Tabe *et al*., 2002). Other than variability caused by nutrient modulation and environmental effects, the relative distribution of seed proteins is primarily genetically determined.

Much of the focus on the proto-industrial production of foreign proteins in seeds has focused on maize, barley and rice, all of which are monocotyledonous seeds that primarily store starch, with oil and protein as less abundant collateral reserve substances (for a review, see Streatfield, 2007). The endosperm of maize, barley and rice has been shown to be a suitable platform capable of producing and sequestering foreign proteins up to a level of 10%–15% of the soluble

protein in seeds, averaging 6%–8% of total protein or approximately 1% of seed weight (for examples, see Hood *et al*., 1997, 2007; Horvath *et al*., 2000; Streatfield *et al*., 2003; Woodard *et al*., 2003; Xue *et al*., 2003; Nandi *et al*., 2005; Takaiwa *et al*., 2007). However, much of the cereal and maize seed protein is in the form of insoluble prolamine protein, yielding foreign protein production levels (when measured as a percentage of soluble protein) of below 1% of the seed mass. For protein biofactory applications, the use of a high-protein dicotyledonous seed, such as soybean, which averages 40% protein, could provide a more optimum conversion of the plant's source nutrient flux for sink conversion to foreign protein products. Dicotyledonous plant model seed platforms, such as tobacco (for examples, see Hoffman *et al*., 1987; Fiedler and Conrad, 1995) and soybean (Moravec *et al*., 2007), have been used to investigate foreign protein accumulation; if the protein is stable, maximum levels of 1%–2% of total protein can be achieved for seeds, 40% protein by weight, thus yielding protein accumulation in the same range as the maximum for maize and rice. Foreign protein gene expression in dicotyledonous seeds has identified certain problems, such as (mis)targeting and post-translational instability caused by protein storage vacuole (PSV)-mediated degradation (Hoffman *et al*., 1988; Pueyo *et al*., 1995; Coleman *et al*., 1996).

To make the best use of the protein synthesis capacity of a high-protein seed, it is desirable to redirect a significant part of the protein synthesis capacity from the production of intrinsic seed proteins to the synthesis of foreign protein(s). Large-scale alteration of the proteome can be produced either by the identification of mutants or the introduction of transgenes that suppress or knockdown one or more of the storage proteins. The exchange of endogenous seed protein for foreign proteins has been investigated as a strategy. In *Arabidopsis*, this has been shown to be an effective strategy to exchange one seed vacuolar storage protein for another (Goossens *et al*., 1999). Although a promising strategy, other attempts using crop platforms, such as rice, have not resulted in a high foreign protein yield (Tada *et al*., 2003). A further complication with the expression of transgenes that suppress intrinsic proteins, or encode the expression of foreign proteins, is that the genetic alteration can result in unpredictable collateral changes in the proteome as the seed rebalances its protein content. Seeds possess intrinsic compositional plasticity resulting from the alteration of the source–sink relationship; this may be perturbed by the accumulation of foreign proteins as an alternative sink protein (for example, see Hagan *et al*., 2003). An unintended change in the proteome caused by collateral rebalancing is one of the more vexing issues of biotechnology, with possible increases in undesirable proteins, such as allergens, among the regulatory concerns that may need to be addressed. Soybeans are a highly productive high-protein crop, whose pattern of storage and protein accumulation in the seeds is well defined (Hill and Breidenbach, 1974; Mienke *et al*., 1981; Herman *et al*., 2003; Hajduch *et al*., 2005). Soybean seeds have been shown to rebalance the proteome when one of the major storage proteins is suppressed. The suppression of the 7S storage protein α/α′ subunit of β-conglycinin (termed βC here) by either transgene or mutation results in the rebalancing of the protein content of soybean seeds with increased 11S glycinin content (Kinney *et al*., 2001; Mori *et al*., 2004); this restores the soybean seed to its standard protein content of approximately 40% (w/w) protein in dry seeds. To develop a protein biofactory platform using the exchange of intrinsic for foreign protein, a predictable outcome is necessary so that each new foreign protein produced does not result in different collateral proteome changes, thereby becoming a regulatory concern. The exchange of glycinin for conglycinin is one example of predictable collateral proteome rebalancing that could be exploited for biotechnology.

A parallel requirement for a protein biofactory platform is to target the transgene product to a stable accumulation site, where it will not be susceptible to degradation or further post-translational modification. Seed storage and other secretory proteins are initially synthesized, processed and assembled in the endoplasmic reticulum (ER) (Vitale and Denecke, 1999). The accretion of newly synthesized proteins in the ER, leading to the formation of protein bodies (PBs), is both a naturally occurring pathway in plants for seed protein accumulation (for a review, see Herman and Larkins, 1999) and a means to accumulate storage proteins and enzyme precursors (for reviews, see Hara-Nishimura *et al*., 1998; Chrispeels and Herman, 2000; Herman and Schmidt, 2004). Foreign protein products can be induced to accrete and form PBs by expressing proteins that naturally accrete in the ER, such as prolamine storage proteins (for examples, see Coleman *et al*., 1996; Kim and Krishnan, 2004), or by adding ER retention sequences that increase the ER residency time (Herman *et al*., 1990), often leading to protein accretion with the formation of PBs (Matsushima *et al*., 2003; Mainieri *et al*., 2004; Moravec *et al*., 2007) in ER subdomains (Torres *et al*., 2001), mimicking the formation of zein and other prolamine PBs (for a review, see Herman and Larkins, 1999). Adding an ER retention sequence to foreign proteins and expressing these proteins in soybean seeds leads to limited (1%–2% w/w) protein accumulation, even when accreted into PBs (Moravec *et al*., 2007). PBs are not normally produced in soybean

seeds, but, in the absence of βC, a portion of the up-regulated glycinin that would normally progress to the PSV is instead sequestered into PBs, which accumulate to a moderately high level and are stable into seed dormancy. The capacity of soybean to produce and retain large numbers of *de novo*-induced PBs of intrinsic proteins, coupled with the increased accumulation of the compensating glycinin, suggests that soybean could prove to be a suitable platform to accumulate a high level of foreign proteins sequestered into PBs.

In this article, it is shown that the introgression of a model foreign protein (GFP-kdel), driven by the glycinin promoter, leads to the formation of PBs in a βC suppression (βCS) background, and results in a fourfold increase in the level of GFP-kdel accumulation to an amount that approaches commercial viability. This result shows that the collateral protein rebalancing mechanism that up-regulates glycinin production in βCS can be exploited to enhance the level of foreign protein production in a soybean production platform.

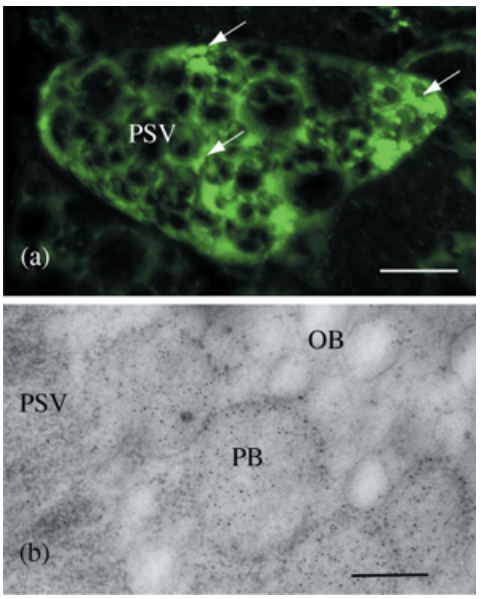

Figure 1 (a) Blue light image of stacked optical sections of a cotyledon storage parenchyma cell of a transgenic soybean seed. Green fluorescent protein (GFP)-kdel is accumulated in the subcellular bodies indicated by arrows. The protein storage vacuoles (PSVs) do not accumulate GFP-kdel and are observed as dark spherical bodies. Bar, 10 μm. (b) Parallel electron microscopic immunoassay using cryofixed tissue labelled with anti-GFP monoclonal antibodies and 5-nm anti-mouse IgG/colloidal gold. The gold particles are observed to be restricted to endoplasmic reticulum (ER)-derived protein bodies (PBs) that possess a limiting membrane with bound ribosomes. OB, oil body. Bar, 1 μm.

Results

Transgenic soybeans produce and sequester GFP-kdel

In order to evaluate the capacity of soybean to accumulate foreign proteins in the same context as proglycinin PBs, an enhanced green fluorescent protein with an in-frame carboxy-terminal ER retention sequence kdel (GFP-kdel), driven by the glycinin promoter, was transformed into soybean. By driving expression with the glycinin seed-specific promoter, the construct was designed to mimic the formation of proglycinin PBs in soybean seed. GFP-kdel has previously been shown to be co-sequestered with other proteins into intrinsic 'ER bodies', a variation of PBs identified in *Arabidopsis* (Matsushima *et al*., 2003). These characteristics make GFP-kdel a good foreign protein model to mimic the proglycinin that is sequestered into PBs formed in βCS soybean (Kinney *et al*., 2001). The expression of GFP-kdel under the control of the seed-specific glycinin promoter resulted in a high level of GFP accumulation at 1.6% of the total protein content. Light microscopic examination of GFP-kdel-expressing cotyledon storage parenchyma cells using a Zeiss multiphonics microscope (Zeiss, Jena, Germany) resulted in the observation of GFP in both diffuse cytoplasmic structures and punctate localization sites. The optical sections comprising a stack were assembled to visualize the three-dimensional distribution of GFP shown in Figure 1a. High-resolution imaging of the GFP-kdel-sequestering organelles was obtained by high-pressure cryofixation of mid-maturation soybean cotyledon tissue, followed by freeze substitution and embedding in Unicryl plastic without the use of any fixative, to optimize tissue samples for immunoreactivity. Thin sections were labelled with anti-GFP monoclonal antibody and indirectly labelled with anti-mouse immunoglobulin G (IgG)/5-nm colloidal gold, and examined and imaged using an electron microscope. The colloidal gold label decorated ER-derived PBs containing a diffuse labelled matrix and bounded by a membrane with attached ribosomes (Figure 1b). Adjacent PSVs were not labelled. The labelling of PBs and the absence of labelling of PSVs correlate with the distribution of GFP fluorescence shown in Figure 1a. The two-dimensional gel GFP spots were identified by anti-GFP antibody cross-reactivity (an example is shown later in Figure 3d), and GFP was quantified as 1.6% protein by spot volume, normalizing to the global protein content (Figure 3b, see later).

Introgression of GFP-kdel into the βCS background results in an increase in GFP level

The increased proglycinin/glycinin accumulated to compensate for the protein shortfall in βCS transgenic (Kinney *et al*., 2001) and mutant (Mori *et al*., 2004) soybeans suggested

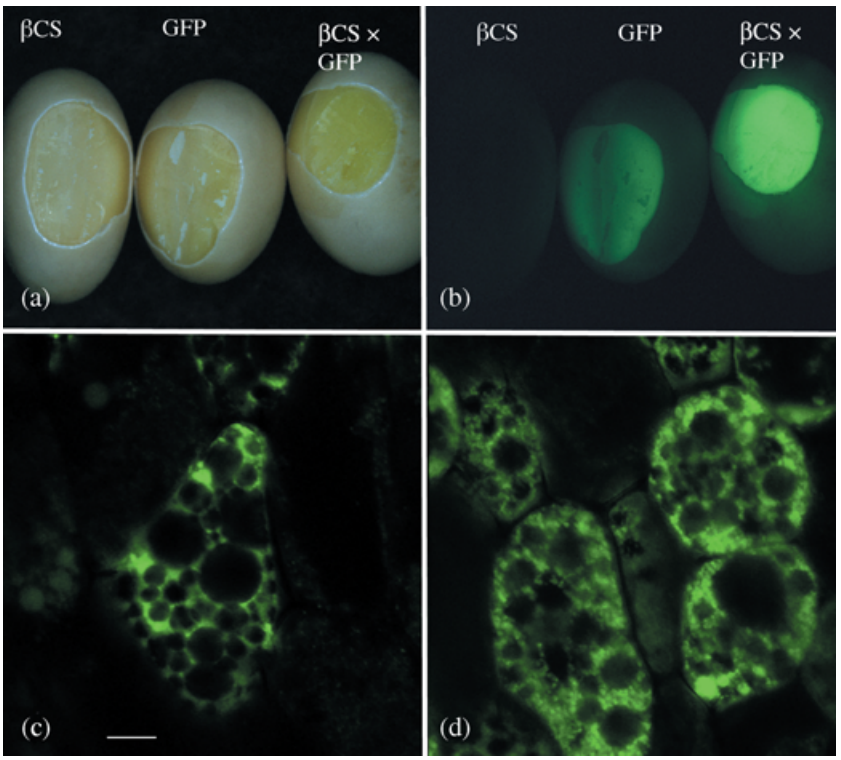

Figure 2 White (a) and blue (b) light images of whole soybean seeds of the two homozygous parental lines [homozygous lines of both green fluorescent protein (GFP)-kdel and α/α′ β-conglycinin suppression (βCS)] and the homozygous progeny of the cross. The seeds shown have been chipped to expose the cotyledon tissue. GFP-kdel and the βCS × GFP-kdel seeds, with identical illumination intensity and imaging exposure, show the differential accumulation of GFP in the βCS × GFP-kdel seeds (d) relative to the GFP-kdel parent (c). The GFP bodies in βCS × GFP-kdel seed cells (d) are observed to have a broader size range, including much larger organelles, than those in the parental GFP-kdel seed (c). (c, d) Bar, 10 μm.

that foreign proteins expressed in the same context may be similarly increased in abundance. To test this hypothesis, the homozygous GFP-kdel line whose GFP expression was driven by the glycinin promoter, which produces an average of 1.6% GFP sequestered into PBs, was introgressed into transgenic βCS soybean (Kinney *et al*., 2001). The resulting R_0 crosses were visually screened under blue light as mature dry seeds, scanning for changes in GFP abundance. The seeds were then chipped and the βCS trait was assayed by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) immunoblot using βC-specific antibody as a probe. Seeds exhibiting both GFP and βCS traits were germinated and grown into plants. The next two generations were assayed for the presence of both traits to assess homozygosity. R_2 homozygous seeds were collected from independent plants and used to assess GFP accumulation in both replicate crosses and in replicate siblings of a cross. Both βCS and GFP-kdel plants were used as pollen donors and as plants brought forward for selection in subsequent generations for use in GFP expression and accumulation analysis.

Homozygous seeds exhibiting both GFP expression and βCS traits were assayed for the accumulation of GFP. The merging of the GFP-kdel and βCS traits resulted in greatly increased GFP fluorescence observed in the intact and chipped seeds of the cross (Figure 2a,b). Portions of the hydrated chips were visualized by fluorescence microscopy using an equal intensity of illumination, and the resulting images showed that the GFP-containing structures were larger and more abundant in the GFP-kdel × βCS line than in the parent GFP-kdel line (Figure 2c,d).

Glycinin promoter-driven GFP-kdel expression in the βCS background results in a fourfold increase in the seed GFP content

Lysates of seed chips were fractionated by two-dimensional, wide-range, isoelectric focusing (IEF)/SDS-PAGE. Figure 3a–c shows two-dimensional gels of both homozygous parental lines and homozygous GFP-kdel × βCS crosses. The GFP spots identified by immunoblot were more abundant in the GFP-kdel × βCS crosses than in the GFP-kdel parent (Figure 3b,c). The GFP spots were quantified in triplicate gels, and the GFP-kdel βCS crosses were determined to contain 7% GFP, compared with 1.6% GFP in the GFP-kdel control. The introgression of GFP into βCS results in an approximately fourfold increase in GFP accumulation.

In order to examine the impact of one PB-sequestered protein, GFP-kdel, on the other, proglycinin, wild-type, βCS and βCS × GFP extracts were fractionated by one-dimensional SDS-PAGE, stained and scanned as shown in Figure 4. The proglycinin and summed processed glycinin (consisting of glycinin A4, acidic glycinin and basic glycinin) bands were

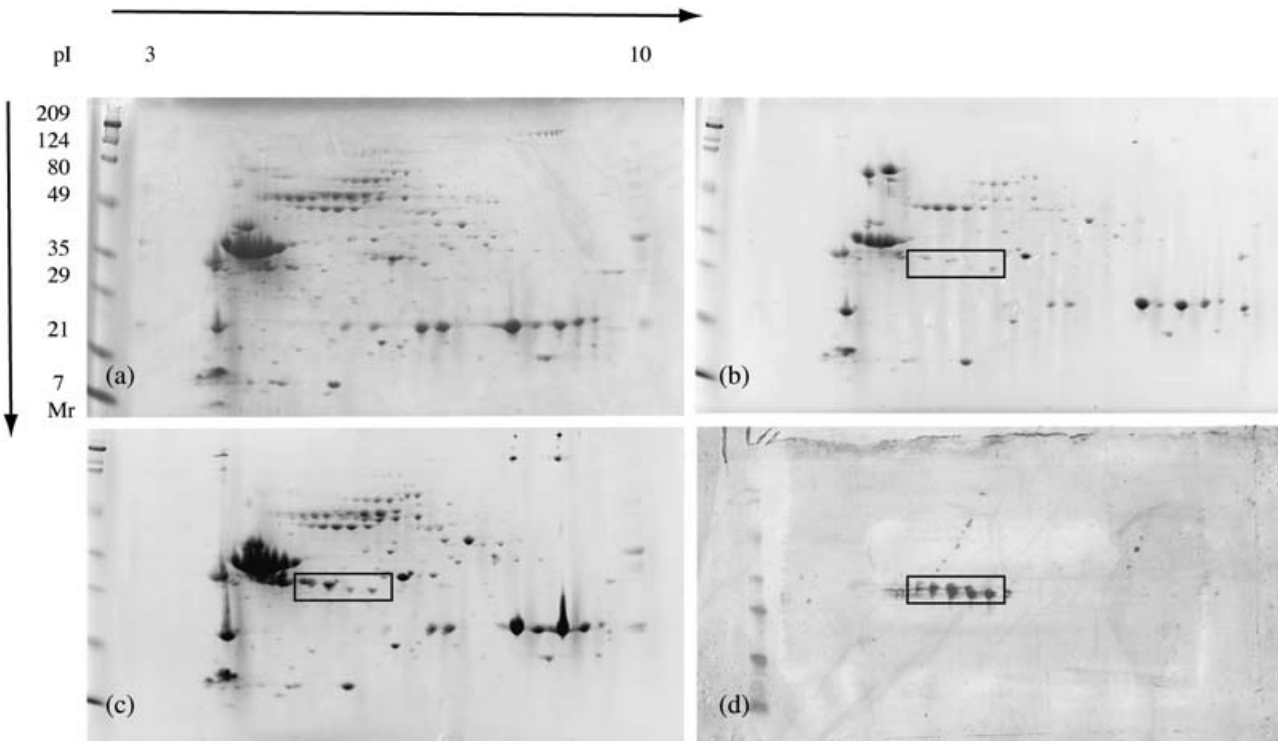

Figure 3 Two-dimensional isoelectric focusing/sodium dodecylsulphate-polyacrylamide gel electrophoresis (IEF/SDS-PAGE) of lysates from homozygous seeds for the two parental lines [α/α′ β-conglycinin suppression (βCS) (a) and green fluorescent protein (GFP)-kdel (b)] and their cross (c). The GFP spots are encircled. Introgression of the GFP-kdel trait into βCS results in enhanced accumulation of GFP-kdel. The increased accumulation of glycinin in βCS is observed by the enhanced abundance of both acidic and basic glycinin subunits in both the parent and the cross. Spot volume measurements of the summed GFP spots, averaged for triplicate gels, resulted in a 1.6% protein average for the GFP-kdel parent and > 7% for the βCS × GFP-kdel cross. The identity of the GFP spots was determined by immunoreactivity on blots using a commercial monoclonal antibody probe (d).

quantified, normalized to 100 and the relative distribution of protein in each of the four glycinin bands was calculated. The results of this assay showed that wild-type proglycinin represented 0% of the total. βCS soybean, which retains a large fraction of its glycinin as PB-sequestered proglycinin (Kinney *et al*., 2001), showed a proglycinin level of 24.8% of the total glycinin fraction. βCS × GFP-kdel soybean, which co-expresses two distinct proteins that accrete into PBs, showed a proglycinin level of 6.95% of the total glycinin fraction. The co-expression of GFP-kdel as a glycinin gene mimic results in more than a threefold decrease in the accretion of proglycinin in the ER and its sequestration into PBs.

To obtain a further evaluation of the GFP abundance, seed lysates were prepared and assayed using a fluorometer with commercial GFP as a control standard. The data obtained included multiple assays of different plants and multiple measurements from each biological replicate. The data were subjected to error analysis and were found to be highly reproducible. The results shown in Figure 5 confirm the observations of both the visual impression of GFP fluorescence (Figure 2) and the spot volume abundance (Figure 3): GFP-kdel × βCS crosses average 3.5–4.0-fold enhancement of GFP accumulation compared with the parent GFP-kdel line. The calculation of GFP accumulation based on a standard curve indicated that the GFP-kdel × βCS seeds contained an average of 28 mg of GFP per gram of dry seed. This corresponds to > 7% GFP as a constituent of the 40% w/w protein content of the soybean seed.

GFP-kdel × βCS soybeans form abundant large PBs that accumulate in the seeds

GFP-kdel has been shown previously to co-accrete as a reporter into an existing population of 'ER bodies' (Matsushima *et al*., 2003) that are analogous to the ER-derived seed storage protein accretions of monocotyledonous seeds (for a review, see Herman and Larkins, 1999), termed 'protein bodies'. There are two types of PB morphology: (i) the protein accretions bud from the ER membrane as discrete organelles; and (ii) the PBs remain contiguous with the ER, as has been documented in maize seeds. Although the term 'protein body' is most often associated with storage protein accretions, it is

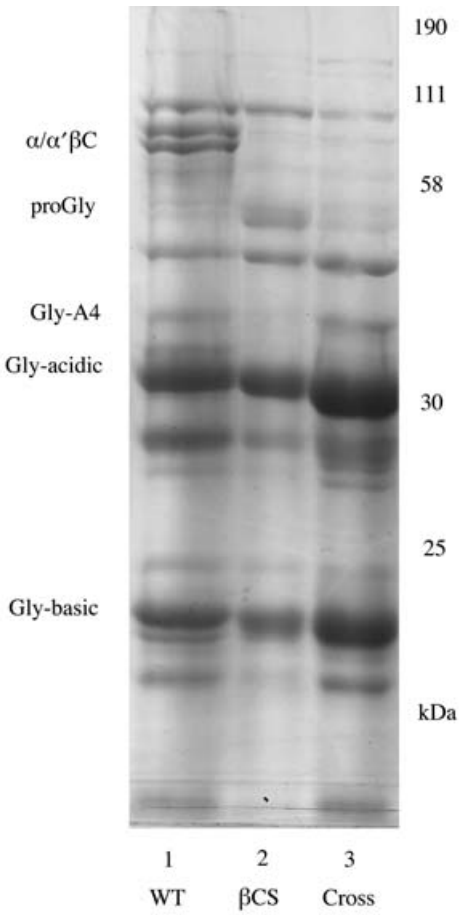

Figure 4 One-dimensional sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS/PAGE)-fractionated seed proteins from cv. Jack (WT), homozygous α/α′ β-conglycinin suppression (βCS) and homozygous βCS × green fluorescent protein (GFP) (Cross) extracts. The resulting stained gels were scanned and the relative distribution of the proglycinin fraction of the summed proglycinin, glycinin A4, glycinin acidic subunit and glycinin basic subunit fractions was determined. The results show a greater than threefold decrease in the proglycinin fraction of the glycinin protein population in the cross.

also used to broadly define any protein accretion within the ER that forms an ER-bounded organelle sequestering a dense protein matrix. In order to obtain more detailed imaging of the GFP-kdel PBs, mid- and late-maturation cotyledon samples were prepared by high-pressure cryofixation, freeze substituted with OsO_4 as a fixative and contrast reagent, and embedded in plastic. Thin sections stained with uranyl acetate were examined by transmission electron microscopy (TEM). Low- to moderate-magnification examination of the cotyledon storage parenchyma cells showed that the cytoplasm of GFP-kdel × βCS soybeans contained abundant novel organelles not observed in the control cv. Jack parent (Figure 6a). The PBs of the GFP-kdel × βCS soybeans, like those of the GFP-kdel soybeans, were bounded by a rough ER-type membrane and sequestered a diffuse matrix of protein with less electron density than the proteins sequestered in the PSV (Figure 6a). Although many of the PBs were observed to be spherical in shape, suggesting that the organelles may have budded from the ER in a similar manner to the PBs found in many grains, including wheat and rice, other PBs appeared to remain interconnected to the ER, as has been shown for maize PBs (Figure 6b). PBs and other ER bodies can be subjected to autophagy. This mechanism transfers the ER-derived PBs to the vacuole in wheat (Levanony *et al*., 1992), and zein PBs assembled in transgenic tobacco seeds are progressively destroyed during seed maturation by autophagy (Coleman *et al*., 1996). The PBs produced by GFP-kdel × βCS seeds did not appear to be subjected to autophagy by the storage parenchyma cell PSVs, which lacked the sequestered, partially disrupted PBs previously observed in maturing tobacco seed (Coleman *et al*., 1996). This is consistent with previous observations of proglycinin PB stability in βCS soybeans (Kinney *et al*., 2001).

Parallel cryofixed GFP × βCS seeds were processed by freeze substitution without added fixative, and embedded in acrylic Lowicryl HM20 resin to provide samples for electron microscope immunocytochemistry. Ultrathin sections were labelled with either anti-GFP mouse monoclonal antibody or polyclonal rabbit anti-glycinin, and then indirectly labelled with anti-species-specific IgG/colloidal gold. As observed in the TEM samples processed for structure, PBs were abundantly distributed in the cytoplasm surrounding the PSVs. The PBs were readily identified by the ribosome-studded limiting membrane, although, in the absence of OsO_4 fixative providing electron density contrast, the matrix of the PBs was observed to be much more electron transparent than the matrix of the adjacent PSVs. Anti-GFP labelled both PBs and ER (Figure 7a), and anti-glycinin densely labelled PSVs and sparsely labelled PBs (Figure 7b). The distribution of the immunolabel showed that the PBs contained dense deposits of GFP and also some glycinin, indicating that the population of PBs obtained by crossing two transgenics (each producing a PB with one protein in its matrix) resulted in the production of a single population of PBs containing both proteins.

Discussion

By mimicking collateral protein rebalancing, the yield of foreign proteins in soybean seeds can be enhanced

The results shown here demonstrate that it is feasible to exploit the collateral proteome rebalancing that occurs in seeds with either gene silencing or the expression of

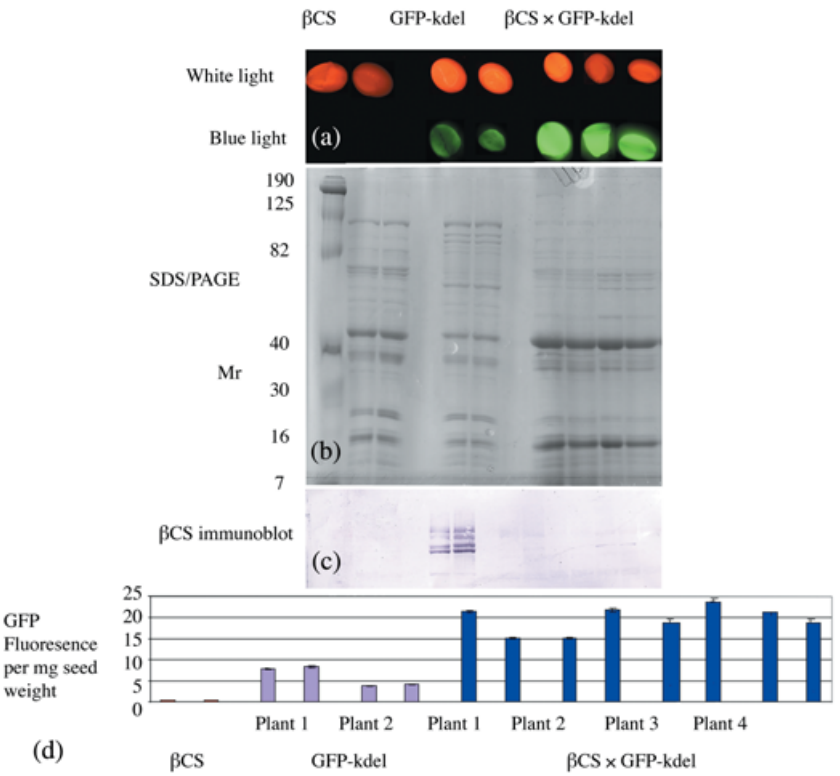

Figure 5 Accumulation of green fluorescent protein (GFP) in the homozygous α/α′ β-conglycinin suppression (βCS) × GFP-kdel cross, compared with the GFP-kdel parent line, determined by fluorometric measurement of seed lysates. (a) White and blue light images of the seeds of the two parental lines and the cross. (b) One-dimensional sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE)-fractionated proteins stained for total protein. (c) Immunoblot using anti-conglycinin antibodies as a probe, demonstrating that only the GFP-kdel parent line possesses β-conglycinin. (d) Graphically displayed data of the GFP content measured by fluorometric assay of the parental lines and the cross. The fluorometric measurements shown demonstrate that the crossed βCS × GFP-kdel seeds possess 3.5–4-fold higher levels of GFP accumulation, to a total of > 7% protein, compared with the 1.6% protein for the parent.

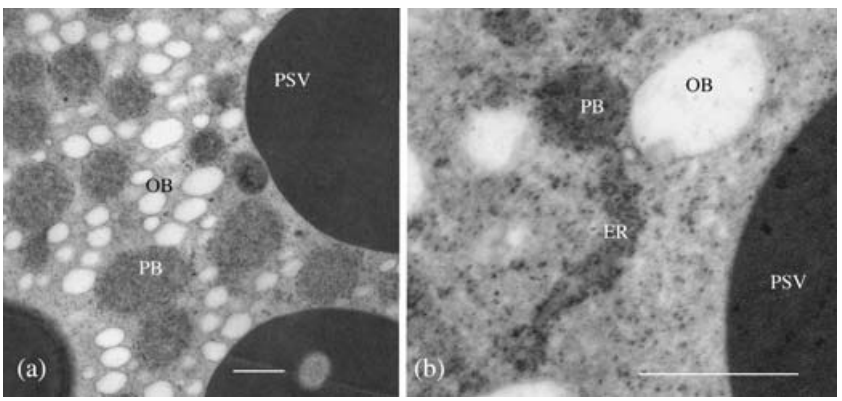

Figure 6 (a) Abundance of protein bodies (PBs) in the cytoplasm of α/α′ β-conglycinin suppression × green fluorescent protein-kdel (βCS × GFP-kdel) late-maturation seed cells. Note that the large numbers of PBs contain a dispersed matrix and are bounded by an endoplasmic reticulum (ER) membrane. The PBs retain the ribosomes which are a characteristic feature of ER-derived PBs. Also shown is an adjacent protein storage vacuole (PSV). (b) The ER origin of PB is illustrated by showing PB attached to a segment of ER. OB, oil body. Bar, 1 μm.

transgenes to produce the enhanced accumulation of foreign proteins. Because collateral changes are not directly correlated to the targeted trait changes, they have the appearance of an unpredictable response with one or more unrelated proteome changes. The key feature of the soybean seed protein production model is that collateral rebalancing appears to be specific, with the GFP-kdel foreign protein sharing the increased accumulation of glycinin protein that results from βCS. The collateral rebalancing of soybean seeds with βCS appears to represent a programmed response to the accumulation shortage of one storage protein (7S βC) by the up-regulation of the accumulation of another major storage protein (11S glycinin), restoring the seed's protein content to the normal wild-type level (Kinney *et al.*, 2001; Mori *et al.*, 2004). Because the rebalancing of the protein content is primarily a result of increased glycinin accumulation, a foreign protein gene mimicking a glycinin gene takes advantage of the soybean's proteome rebalancing process and enhances the yield of the GFP foreign protein. Although tests need to be conducted with several different proteins, it is probable that the rebalancing process may be standardized with a foreign protein substituting for some of the increased glycinin without inducing other collateral proteome alterations. This will have obvious regulatory advantages, because the changes that occur in the soybean proteome may be essentially standardized using this expression strategy. There

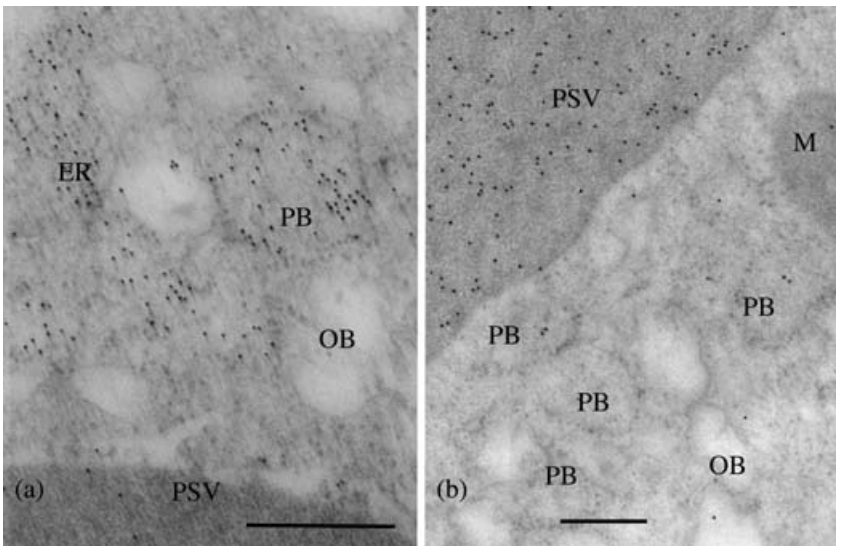

Figure 7 Electron microscopy immunogold assay of the α/α′ β-conglycinin suppression × green fluorescent protein-kdel (βCS × GFP-kdel) seed protein bodies (PBs) using anti-GFP and anti-glycinin. (a) Anti-GFP labels both the endoplasmic reticulum (ER) and PBs, but there is little label seen on the protein storage vacuole (PSV). (b) The anti-glycinin label is localized primarily on PSV with sparse labelling on PB. A mitochondrion (M) and oil body (OB) are also shown. Bar, 1 μm.

is a limitation to this strategy in the expected maximum yield that may be achieved for the foreign protein. By sharing the increase in glycinin accumulation with a foreign protein, it is expected that the maximum accumulation of the foreign protein will be capped at the level of βCS to be replaced. With βCS accounting for 15%–20% of the total soybean seed protein, more than 7% of GFP-kdel accumulated in exchange for βCS represents a large fraction of the total protein rebalancing content potentially available. This approaches an economically viable level of foreign protein production in soybean.

PBs can stably sequester abundant foreign protein in soybean seeds

The addition of ER retention proteins to foreign proteins expressed in plants retards the exit of the protein from the ER (Herman *et al*., 1990; Wandelt *et al*., 1992). ER retention, promoting protein accretion into PBs, is now a widely used strategy to stabilize foreign protein products in plants (for a review, see Vitale and Pedrazzini, 2005). Although PB formation sequesters proteins into an inert cytoplasmic compartment that is an optimum organelle for foreign protein accumulation, the benefit of this approach is often transient in many plant species and tissues. Plant vacuoles, as the lytic compartment in most cells, continually participate in protein turnover and cellular remodelling by autophagy. The lytic character of the vacuole is the primary impediment to its use as a cellular destination of transgene products. PBs produced in wheat seeds and PBs assembled from zeins in transgenic plants are sequestered into the vacuole by autophagy and, for foreign proteins, this results in post-translational instability (Levanony *et al*., 1992; Coleman *et al*., 1996). Experiments on transgenic tobacco seeds have shown that, although PSVs stably accumulate intrinsic storage proteins, the proteins produced by transgenes, whether by targeting to PSVs by endomembrane progression (Pueyo *et al*., 1995) or secondarily accumulated in PSVs by autophagy (Coleman *et al*., 1996), are post-translationally unstable. In tobacco seeds, although a high level of foreign proteins can be synthesized, the subsequent vacuole-mediated degradation leaves little of this protein remaining in the dormant seeds. The discovery that soybeans produce proglycinin PBs with βCS, and that these PBs are stable through seed maturation into dormancy (Kinney *et al*., 2001) and are still present in the cells when the seed germinates, shows that soybeans function differently from tobacco seeds (Coleman *et al*., 1996), and this difference has potential biotechnological applications. The stability of PBs sequestering foreign proteins, accreted with the aid of ER retention sequences in soybean seeds, has been demonstrated by the production of both GFP-kdel and *Escherichia coli* heat-labile toxin B subunit (Moravec *et al*., 2007).

The *de novo*-formed PBs in soybean seeds are large and sequester multiple proteins relative to the PBs normally produced in plants

Protein accretion in the ER, leading to the formation of PBs and similar organelles sequestering storage proteins and precursor proteins, has been described in many plants. Although PB formation appears to be triggered by the accretion of the proteins sequestered, there still remain many questions about whether this is the result of the physical properties of the proteins, such as with the hydrophobic prolamine proteins, or whether it is facilitated by ER-resident proteins

and conditions. The PBs formed by βCS × GFP-kdel soybean are notable with regard to two characteristics: their size and apparent low electron density protein matrix. In appearance, the matrix of βCS × GFP-kdel PBs differs from that of βCS proglycinin PBs, which exhibit a densely packed matrix (Kinney *et al*., 2001). It is interesting that the co-expression of GFP-kdel in the βCS background reduces the glycinin retained in its PB precursor form by at least threefold. This indicates that the presence and accretion of GFP-kdel somehow impedes the accretion of proglycinin, permitting a larger fraction to progress through the endomembrane system to the vacuole, where it is processed into acidic and basic subunits. It has been assumed that the accretion of proteins occurs via their chemical properties, and that this may be modified by the co-expression of another accreting protein. Alternatively, if protein accretion is an active process requiring the participation of one or more ER lumen proteins, the co-expression of GFP-kdel with increased glycinin, leading to a decrease in the amount of glycinin retained in its precursor form, may result from a greater affinity of GFP-kdel to one or more ER lumen proteins that promote the accretion of proteins into PBs.

The intrinsic PBs of cereals and maize are primarily relatively small organelles, as are the PBs sequestering enzyme and storage protein precursors. The PBs formed from the expression of foreign proteins are primarily ER-bounded structures sequestering a dense protein matrix. The βCS × GFP-kdel PBs accumulated in soybeans maintain a normal PB morphology of a spherical organelle bounded by an ER-derived membrane which may have continuity with the ER network. βCS × GFP-kdel PBs are larger organelles than those typically observed in plants, and the matrix of PBs containing both GFP-kdel and glycinin is less densely packed than that of comparable PBs, including those formed in soybean containing proglycinin (Kinney *et al*., 2001), *E. coli* heat-labile toxin B subunit (Moravec *et al*., 2007) and zein (Kim and Krishnan, 2004). The size, protein packing and stability of PBs in soybean seeds show that soybean seeds tolerate the engineering of a large fraction of the cytoplasmic space to be occupied by PBs maintained as stable organelles through seed maturation.

Soybean seeds represent a potentially useful platform as a protein bioreactor

The results presented here indicate that it is possible to increase the yield of foreign protein accumulation from 1.6% to approximately 7.0% by trading the capacity for intrinsic protein synthesis for the capacity to produce foreign proteins. By linking the enhancement of sequestration of a foreign protein in ER-derived PBs to the collateral protein compensation mechanism, the quantity of foreign protein accumulated in soybean seeds can be increased significantly. As a high-protein platform, soybean could prove to be an ideal tool for commercial-scale foreign protein production. The potential applications stemming from the enhanced foreign protein yield of a high-protein bioreactor platform, such as soybean, include many moderate-value protein products. These applications include enzymes and biocatalysts, where the scale, low-cost production, and ease of shipping and processing conferred using soybean seeds can be employed to an advantage.

Experimental procedures

Transgenic soybean seeds

The parameters for soybean (*Glycine max* L. Merrill cv. Jack) somatic embryo transformation by biolistics have been described previously (Trick *et al*., 1997). Regeneration was performed as described in Schmidt *et al*. (2004). The hygromycin resistance gene (kindly provided by N. Murai, Louisiana State University, Baton Rouge, LA, USA), under the control of potato ubiquitin 3 regulatory elements (Garbarino and Belknap, 1994), was used as a selectable marker in tissue culture. A commercially available GFP (Clontech Inc., Mountain View, CA, USA) open reading frame was placed into a cassette containing the seed-specific glycinin regulatory elements (Nielsen *et al*., 1989), the 20-amino-acid ER signal sequence from the *Arabidopsis* chitinase gene and kdel retention tag, as described previously (Moravec *et al*., 2007). Mature dry seeds were harvested and visually observed for GFP expression under a fluorescent dissecting microscope using blue (450-nm) light for excitation. GFP-positive plants were grown to the T_2 generation to obtain homozygous seeds. GFP seeds were examined at the cellular level using a two-photon excitation Zeiss LSM 510 microscope, with excitation at 488 nm and a 512-nm emission filter.

Two-dimensional protein analysis

Total protein was isolated from mature soybean seeds as described previously (Joseph *et al*., 2004), and a total of 150 μg protein was loaded on to an 11-cm immobilized pH gradient (IPG) gel strip (pH 3–10 non linear) (Bio-Rad, Hercules, CA, USA) and hydrated overnight. IEF was performed for a total of 40 kVh using a Protean IEF Cell (Bio-Rad), and then run in the second dimension SDS-PAGE gel (8%–16% linear gradient). Gels were stained overnight in 0.1% Coomassie blue in 40% (v/v) methanol and 10% (v/v) acetic acid. Blotting and subsequent immunodetection using GFP monoclonal antibody (Clontech Inc.) were performed as described previously (Joseph *et al*., 2004). Each sample was run on triplicate gels, and scanned and analysed on a Phoretix 2D Evolution (version 2005; Nonlinear Dynamics Ltd., Durham, NC, USA). The GFP spots identified on the immunoblot allowed the corresponding spots to be located on the replicate gels, and the volumes of these spots were normalized against the entire proteome spot volume to determine the percentage volume of GFP in the entire soluble soybean seed proteome.

α/α′ βC knockdown cross GFP seeds

Once homozygous GFP-kdel plants had been detected by visual fluorescence and homozygous βCS plants had been observed by one-dimensional SDS-PAGE gels, the plants were cross-pollinated. GFP plants were used as the pollen source so that putative cross seeds were readily identifiable by their fluorescence under blue light. Crossed seeds were then chipped and assayed by one-dimensional SDS-PAGE immunoblot analysis using polyclonal anti-βC antibodies to confirm the βCS phenotype. Seeds with the βCS phenotype and GFP-kdel phenotype were planted and grown to homozygous generation.

Quantification of GFP

Chipped seeds were weighed (approximately 10 mg), and protein was extracted in 10 mL of extraction buffer [0.1 M tris(hydroxymethyl)aminomethane (Tris), pH 8; 0.1% Triton]. Fluorescence was read on a fluorescence spectrophotometer (Varian Inc., Palo Alto, CA, USA) with excitation at 488 nm and emission at 510 nm. Recombinant GFP (Vector Laboratories, Burlington, CA, USA) was used to produce a standard curve. The amount of fluorescence in the soybean seeds was compared with this standard curve to calculate GFP units/mg protein. Buffer was used as a blank to zero the fluorescence on the spectrophotometer. Seed chips from the two parental lines (βCS and GFP-kdel) and from four homozygous cross lines were used to assay the level of GFP. Two plants were assayed for each line, and each sample was read in triplicate with the average reading ± standard error recorded.

Electron microscopy and immunogold immunocytochemistry

Tissue samples were cryofixed with a Balzer's high-pressure device (Bal-Tech, Principality of Liechtenstein), freeze substituted with acetone–OsO_4 and embedded in epon plastic. Ultrathin sections were stained with both saturated aqueous uranyl acetate and lead citrate (33 mg/mL) prior to observation. For immunocytochemical analysis, parallel samples were cryofixed and processed by freeze substitution without any fixative. The substituted samples were transferred to Lowicryl HM-20 resin, which was polymerized by UV light illumination. Thin sections were labelled with anti-GFP monoclonal antibody (Clontech Inc.) or rabbit polyclonal anti-glycinin previously produced by this laboratory. The sections were indirectly labelled with anti-IgG (rabbit or mouse) coupled to either 5 nm or 10 nm colloidal gold (Sigma, St. Louis, MO, USA), and then contrasted with 5% uranyl acetate before electron microscopy observation. TEM was performed with a LEO 912AB microscope, with images captured using a 2k × 2k charge-coupled device (CCD) camera operated in the montage mode.

Acknowledgements

We thank Drs Howard Berg and Leslie Hicks (Donald Danforth Plant Science Center) for their assistance with microscopy and proteomics assays. We thank Dr Anthony Kinney (Pioneer Hybrid/Dupont, Wilmington, DE, USA) for the α/α′ subunit of βC-suppressed seeds described in Kinney *et al*. (2001).

References

Beach, L.R., Spencer, D., Randall, P.J. and Higgins, T.J.V. (1985) Transcriptional and post-transcriptional regulation of storage protein gene expression in sulfur-deficient pea seeds. *Nucleic Acids Res*. **13**, 999–1013.

Chrispeels, M.J. and Herman, E.M. (2000) UPDATE-ER-derived compartments function in storage and as mediators of vacuolar remodeling via a new type of organelle, precursor protease vesicles (PPV). *Plant Physiol*. **123**, 1227–1233.

Coleman, C.E., Herman, E.M., Takasaki, K. and Larkins, B.A. (1996) The maize γ-zein sequesters α-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. *Plant Cell*, **8**, 2335–2345.

Fiedler, U. and Conrad, U. (1995) High-level production and long-term storage of engineered antibodies in transgenic tobacco seeds. *Biotechnology (NY)*, **13**, 1090–1093.

Garbarino, J.E. and Belknap, W.R. (1994) Isolation of a ubiquitin-ribosomal protein gene (Ubi3) from potato and expression of its promoter in transgenic plants. *Plant Mol. Biol*. **24**, 119–127.

Goossens, A., Van Montagu, M. and Angenon, G. (1999) Co-introduction of an antisense gene for an endogenous seed storage protein can increase expression of a transgene in *Arabidopsis thaliana* seeds. *FEBS Lett*. **456**, 160–164.

Hagan, N.D., Upadhyaya, N., Tabe, L.M. and Higgins, T.J. (2003) The redistribution of protein sulfur in transgenic rice expressing a gene for a foreign, sulfur-rich protein. *Plant J*. **34**, 1–11.

Hajduch, M., Ganapathy, A., Stein, J.W. and Thelen, J.J. (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol*. **137**, 1397–1419.

Hara-Nishimura, I., Shimada, T., Hatano, K., Takeuchi, Y. and Nishimura, M. (1998) Transport of storage proteins to protein storage vacuoles is mediated by large precursor-accumulating vesicles. *Plant Cell*, **10**, 825–836.

Herman, E.M. and Larkins, B.A. (1999) Protein storage bodies. *Plant Cell*, **11**, 601–661.

Herman, E.M. and Schmidt, M.A. (2004) Endoplasmic reticulum to vacuole trafficking of endoplasmic reticulum bodies provides an alternate pathway for protein transfer to the vacuole. *Plant Physiol*. **136**, 3440–3446.

Herman, E.M., Tague, B.W., Hoffman, L.M., Kjemtrup, S.E. and Chrispeels, M.J. (1990) Retention of phytohemagglutinin with carboxyterminal tetrapeptide KDEL in the nuclear envelope and the endoplasmic reticulum. *Planta*, **182**, 305–312.

Herman, E.M., Helm, R.M., Jung, R. and Kinney. A.J. (2003) Genetic modification removes an immunodominant allergen from soybean. *Plant Physiol*. **132**, 36–43.

Hill, J.E. and Breidenbach, R.W. (1974) Proteins of soybean seeds. II. Accumulation of the major protein components during seed development and maturation. *Plant Physiol*. **53**, 747–751.

Hirai, M.Y., Fujiwara, T., Chino, M. and Naito, S. (1995) Effects of sulfate concentrations on the expression of a soybean seed storage protein gene and its reversibility in transgenic *Arabidopsis thaliana*. *Plant Cell Physiol*. **36**, 1331–1339.

Hoffman, L.M., Donaldson, D.D., Bookland, R., Rashka, K. and Herman, E.M. (1987) Synthesis and protein body deposition of maize 15-kd zein in transgenic tobacco seeds. *EMBO J*. **6**, 3213–3221.

Hoffman, L.M., Donaldson, D.D. and Herman, E.M. (1988) A modified storage protein is synthesized, processed, and degraded in seeds of transgenic plants. *Plant Mol. Biol*. **11**, 717–729.

Hood, E.E., Witcher, D.R., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R.J., Hernan, R., Kappel, W.K., Ritland, D., Li, C.-P. and Howard, J.A. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Mol. Breed*. **3**, 291–306.

Hood, E.E., Lane, R.L., Bray, J., Clough, R., Pappu, K., Drees, C., Hood, K.R., Yoon, S., Ahmad, A. and Howard, J.A. (2007) Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed. *Plant Biotechnol. J*. **5**, 709–719.

Horvath, H., Huang, J., Wong, O., Kohl, E., Okita, T., Kannangara, C.G. and von Wettstein, D. (2000) The production of recombinant proteins in transgenic barley grains. *Proc. Natl. Acad. Sci. USA*, **97**, 1914–1919.

Joseph, L.M., Hymowitz, T., Schmidt, M.A. and Herman, E.M. (2004) Evaluation of *Glycine* germplasm for nulls of the immunodominant allergen P34/Gly m Bd 30k. *Crop Sci*. **46**, 1755–1763.

Kim, W.S. and Krishnan, H.B. (2004) Expression of an 11 kDa methionine-rich delta-zein in transgenic soybean results in the formation of two types of novel protein bodies in transitional cells situated between the vascular tissue and storage parenchyma cells. *Plant Biotechnol. J*. **2**, 199–210.

Kinney, A.J., Jung, R. and Herman, E.M. (2001) Cosuppression of the α-subunits of β-conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein bodies. *Plant Cell*, **13**, 1165–1178.

Levanony, H., Rubin, R., Altschuler, Y. and Galili, G. (1992) Evidence for a novel route of wheat storage proteins to vacuoles. *J. Cell Biol*. **119**, 1117–1128.

Mainieri, D., Rossi, M., Archinti, M., Bellucci, M., De Marchis, F., Vavassori, S., Pompa, A., Arcioni, S. and Vitale, A. (2004) Zeolin. A new recombinant storage protein constructed using maize zein and bean phaseolin. *Plant Physiol*. **136**, 3447–3456.

Matsushima, R., Hayashi, Y., Yamada, K., Shimada, T., Nishimura, M. and Hara-Nishimura, I. (2003) The ER body, a novel endoplasmic reticulum-derived structure in Arabidopsis. *Plant Cell Physiol*. **44**, 661–666.

Mienke, D.W., Chen, J. and Beachy, R.N. (1981) Expression of storage-protein genes during soybean seed development. *Planta*, **153**, 130–139.

Moravec, T., Schmidt, M.A., Herman, E.M. and Woodford-Thomas, T. (2007) Production of *Escherichia coli* heat labile toxin (LT) B subunit in soybean seed and analysis of its immunogenicity as an oral vaccine. *Vaccine*, **25**, 1647–1657.

Mori, T., Maruyama, N., Nishizawa, K., Higasa, T., Yagasaki, K., Ishimoto, M. and Utsumi, S. (2004) The composition of newly synthesized proteins in the endoplasmic reticulum determines the transport pathways of soybean seed storage proteins. *Plant J*. **40**, 238–249.

Nandi, S., Yalda, D., Lu, S., Nikolov, Z., Misaki, R., Fujiyama, K. and Huang, N. (2005) Process development and economic evaluation of recombinant human lactoferrin expressed in rice grain. *Transgenic Res*. **14**, 237–249.

Nielsen, N.C., Dickinson, C.D., Cho, T.J., Thanh, V.H., Scallon, B.J., Fischer, R.L., Thomas, L.S., Drews, G.N. and Goldberg, R.B. (1989) Characterization of the glycinin gene family in soybean. *Plant Cell*, **1**, 313–328.

Pueyo, J.J., Chrispeels, M.J. and Herman, E.M. (1995) Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole. *Planta*, **196**, 586–596.

Schmidt, M.A., Tucker, D.M., Cahoon, E.B. and Parrott, W.A. (2004) Towards normalization of soybean somatic embryo maturation. *Plant Cell Rep*. **24**, 383–391.

Streatfield, S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol. J*. **5**, 2–15.

Streatfield, S.J., Lane, J.R., Brooks, C.A., Barker, D.K., Poage, M.L., Mayor, J.M., Lamphear, B.J., Drees, C.F., Jilka, J.M., Hood, E.E. and Howard, J.A. (2003) Corn as a production system for human and animal vaccines. *Vaccine*, **21**, 812–815.

Tabe, L., Hagan, N. and Higgins, T.J. (2002) Plasticity of seed protein composition in response to nitrogen and sulfur availability. *Curr. Opin. Plant Biol*. **5**, 212–217.

Tada, Y., Utsumi, S. and Takaiwa, F. (2003) Foreign gene products can be enhanced by introduction into low storage protein mutants. *Plant Biotechnol. J*. **1**, 411–422.

Takaiwa, F., Takagi, H., Hirose, S. and Wakasa, Y. (2007) Endosperm tissue is good production platform for artificial recombinant proteins in transgenic rice. *Plant Biotechnol. J*. **5**, 84–92.

Torres, E., Gonzalez-Melendi, P., Stöger, E., Shaw, P., Twyman, R.M., Nicholson, L., Vaquero, C., Fischer, R., Christou, P. and Perrin, Y. (2001) Native and artificial reticuloplasmins co-accumulate in distinct domains of the endoplasmic reticulum and in post-endoplasmic reticulum compartments. *Plant Physiol*. **127**, 1212–1223.

Trick, H.N., Dinkins, R.D., Di Santarem, E.R.R., Samoylov, V., Meurer, C.A., Walker, D., Parrott, W.A., Finer, J.J. and Collins, G.B. (1997) Recent advances in soybean transformation. *Plant Tiss. Cult. Biotechnol*. **3**, 9–26.

Vitale, A. and Denecke, J. (1999) The endoplasmic reticulum: gateway of the secretory pathway. *Plant Cell*, **11**, 615–662.

Vitale, A. and Pedrazzini, E. (2005) Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. *Mol. Intervent*. **5**, 216–225.

Wandelt, C.I., Khan, M.R.I., Craig, S., Schroeder, H.E., Spencer, D. and Higgins, T.J.V. (1992) Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *Plant J*. **2**, 181–192.

Woodard, S.L., Mayor, J.M. and Bailey, M.R. (2003) Maize (*Zea mays*)-derived bovine trypsin: characterization of the first large-scale, commercial protein product from transgenic plants. *Biotechnol. Appl. Biochem*. **38**, 123–130.

Xue, G.P., Patel, M., Johnson, J.S., Smyth, D.J. and Vickers, C.E. (2003) Selectable marker-free transgenic barley producing a high level of cellulase (1,4-beta-glucanase) in developing grains. *Plant Cell Rep*. **21**, 1088–1094.